



OPEN ACCESS

EDITED BY

Tianzhen Zhang,
Zhejiang University, China

REVIEWED BY

Shengjun Li,
Qingdao Institute of Bioenergy and
Bioprocess Technology, Chinese
Academy of Sciences (CAS)
N.C. Ereful,
University of the Philippines Los
Banos, Philippines

*CORRESPONDENCE

Lirui Cheng

✉ chenglirui@caas.cn

Aiguo Yang

✉ yangaiguo@caas.cn

Yuanying Wang

✉ wangyuanying@caas.cn

SPECIALTY SECTION

This article was submitted to
Plant Breeding,
a section of the journal
Frontiers in Plant Science

RECEIVED 01 November 2022

ACCEPTED 02 December 2022

PUBLISHED 10 January 2023

CITATION

Yuan G, Sun K, Yu W, Jiang Z, Jiang C,
Liu D, Wen L, Si H, Wu F, Meng H,
Cheng L, Yang A and Wang Y (2023)
Development of a MAGIC population
and high-resolution quantitative
trait mapping for nicotine
content in tobacco.
Front. Plant Sci. 13:1086950.
doi: 10.3389/fpls.2022.1086950

COPYRIGHT

© 2023 Yuan, Sun, Yu, Jiang, Jiang, Liu,
Wen, Si, Wu, Meng, Cheng, Yang and
Wang. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Development of a MAGIC population and high-resolution quantitative trait mapping for nicotine content in tobacco

Guangdi Yuan, Kefan Sun, Wenlong Yu, Zipeng Jiang,
Caihong Jiang, Dan Liu, Liuying Wen, Huan Si, Fengyan Wu,
He Meng, Lirui Cheng*, Aiguo Yang* and Yuanying Wang*

Key Laboratory of Tobacco Improvement and Biotechnology, Tobacco Research Institute, Chinese Academy of Agricultural Sciences, Qingdao, China

Multiparent Advanced Generation Inter-Cross (MAGIC) population is an ideal genetic and breeding material for quantitative trait locus (QTL) mapping and molecular breeding. In this study, a MAGIC population derived from eight tobacco parents was developed. Eight parents and 560 homozygous lines were genotyped by a 430K single-nucleotide polymorphism (SNP) chip assay and phenotyped for nicotine content under different conditions. Four QTLs associated with nicotine content were detected by genome-wide association mapping (GWAS), and one major QTL, named *qNIC7-1*, was mapped repeatedly under different conditions. Furthermore, by combining forward mapping, bioinformatics analysis and gene editing, we identified an ethylene response factor (ERF) transcription factor as a candidate gene underlying the major QTL *qNIC7-1* for nicotine content in tobacco. A presence/absence variation (PAV) at *qNIC7-1* confers changes in nicotine content. Overall, the large size of this MAGIC population, diverse genetic composition, balanced parental contributions and high levels of recombination all contribute to its value as a genetic and breeding resource. The application of the tobacco MAGIC population for QTL mapping and detecting rare allelic variation was demonstrated using nicotine content as a proof of principle.

KEYWORDS

tobacco, MAGIC population, GWAS, nicotine content, QTL

Introduction

Most traits of biological and economic interest in crop plants are under polygenic control, displaying complicated genetic structure and are extensively influenced by the interaction of genotype and environment (Falconer and Mackay, 1996; Doerge, 2002; Holland, 2007). Conventionally, the identification of genes underlying agronomic

quantitative trait loci (QTLs) is performed using biparental mapping populations, such as recombinant inbred lines (RILs), and double haploid (DH) population (Rakshit et al., 2012). However, these populations have limitations due to the narrow genetic variation and limited opportunities for genetic recombination events, which lead to a lower map resolution (Valdar et al., 2006; Chen et al., 2013; Dell'Acqua et al., 2015; Huang et al., 2015).

Genome-wide association study (GWAS) provides another strategy for discovering genes and regions associated with agronomic traits using diverse population sets (Mitchell-Olds et al., 2010; Ongom and Ejeta, 2018). GWAS using SNP markers have made outstanding achievements in many economically important crops due to the exploitation of a large number of historical recombination events that lead to the rapid decay of linkage disequilibrium (Remington et al., 2001; Jaiswal et al., 2019). The mapping resources that have been widely employed for GWAS are accessions of landrace genetic resources or breeding lines that have been extensively phenotyped (Ongom and Ejeta, 2018). However, the main limitations of GWAS are linkage disequilibrium (LD), population substructure, and unbalanced allele frequencies (Mitchell-Olds, 2010; Visscher et al., 2012).

The main drawbacks of biparental populations and germplasm collections may be addressed by Multiparent Advanced Generation Inter-Cross (MAGIC) population (Mackay and Powell, 2007; Pascual et al., 2015; Scott et al., 2020). MAGIC population is a fine-scale mosaics of RILs that have roughly equal proportions of the founder genomes (Mackay and Powell, 2007; Cavanagh et al., 2008; Rakshit et al., 2012). The multiple founders of MAGIC population can enrich allelic diversity, whereas multiple intercrossing cycles result in a set of rearranged genomes with a high level of fragmentation, which gives greater opportunity for recombination and dramatically increases the power for QTL detection (Chen et al., 2013; Yamamoto et al., 2014; Arrones et al., 2020). Additionally, the restrictions of population structure and rare alleles for GWAS were mitigated to some extent in MAGIC population. Therefore, MAGIC population provides a compromise between the much greater complexity existed in naturally occurring accessions and the extreme simplicity of a diallelic system of RILs, making MAGIC an ideal material for GWAS analysis (Kover et al., 2009). To date, MAGIC population has been widely used in QTL mapping in many crop species. Bossa-Castro et al. (2018) identified 11 BSR QTLs effective against both bacterial leaf streak and bacterial blight and 3 pathovar-specific QTLs using an eight-way rice MAGIC population. Thyssen et al. (2019) identified genomic loci and candidate genes for six major fiber quality traits in a cotton MAGIC population using GWAS and whole genome sequencing. Rollar et al. (2021) used a Bavarian MAGIC wheat population for simple interval mapping and identified 19 QTLs corresponding to 11 distinct chromosomal regions controlling

leaf rust resistance. However, to date, the construction and QTL mapping of MAGIC population have not been reported in tobacco.

Tobacco (*Nicotiana tabacum* L.) is an economically important crop and is also a model for the investigation of plant pathology, genetics, and biotechnology. Studies of tobacco genetics and genomics strengthened the use of *N. tabacum* as a plant model system and were also beneficial for tobacco molecular breeding. In tobacco, nicotine is an abundant predominant alkaloid produced exclusively in the roots and accumulating mainly in the leaves. It is derived from diamine putrescine, which depends on a series of metabolic enzymes, including putrescine N-methyltransferase (PMT), quinolinate phosphoribosyl transferase (QPT) and the PIP-family oxidoreductase A622 (Hibi et al., 1992; Sinclair et al., 2000; Kajikawa et al., 2009; Ryan et al., 2012). The biosynthesis of nicotine is genetically controlled by two distinct loci, NICOTINE1 (*NIC1*) and *NIC2* (Legg and Collins, 1971; Hibi et al., 1994). It has been reported that the clustered ethylene response factor (ERF) transcription factor at the *NIC2* locus can regulate nicotine content in tobacco by regulating PMT and QPT expression (Shoji et al., 2010; Qin et al., 2021). Sun et al. (2017) reported that nicotine content exhibited extreme variation in different tobacco accessions, and two significant SNP markers associated with nicotine content were identified in tobacco using the GWAS method. In this study, we developed the first eight-way MAGIC population in tobacco and identified genes/QTLs associated with nicotine contents using the MAGIC population. Based on 93,766 SNP markers and the phenotypic data obtained from three environments, four quantitative trait loci (QTLs) associated with nicotine content trait were mapped on Chr07 and Chr05 by GWAS. By combining genetic population research, bioinformatics analysis and gene editing, an ERF transcription factor was considered as the candidate gene underlying the major QTL *qNIC7-1* for nicotine content in tobacco. A presence/absence variation (PAV) at *qNIC7-1* confers changes in nicotine content. These results demonstrated the advantage and confidence of MAGIC population in the high-resolution detection of QTLs and genes responsible for complex traits.

Materials and methods

Tobacco MAGIC Population

A total of 8 tobacco accessions were used as parents of the MAGIC population. The construction of the MAGIC population used in this study is shown in Figure 1A. Beinhart1000-1 (BH) and Florida301 are cigar tobacco. Basma and Samsun are oriental tobacco. Xiaohuaqing (XHQ) and Tangpeng (TP) are sun-cured tobacco from China. Vam belongs to burley tobacco. Honghuadajinyuan (HD) is a flue-cured tobacco cultivar. The

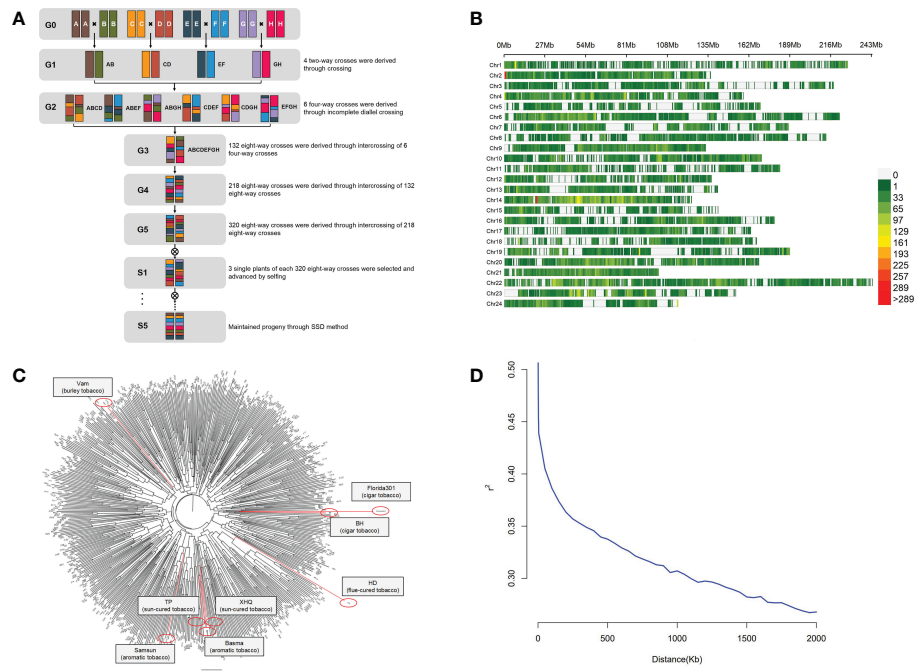


FIGURE 1 Diversity in the tobacco MAGIC population. **(A)** Breeding scheme of tobacco MAGIC population. Individual color represents each of eight parents (BH, Florida301, Vam, Basma, Samsun, XHQ, TP, HD). **(B)** Distribution of SNPs in 24 chromosomes. **(C)** Neighbor-joining tree of the 8 parents and 560 MAGIC lines, with parents labeled in red. **(D)** Decay of linkage-disequilibrium (LD).

first stage was pairwise crossing of eight parents, and 4 two-way crosses were produced. The resulting 4 two-way crosses were incomplete diallel crossed to produce 6 four-way crosses. The 6 crosses were then intercrossed to produce 132 eight-way crosses. To increase the number of recombination events, the resulting eight-way crosses were intercrossed to produce 218 eight-way crosses, and the 218 crosses were randomly mated and produced 320 eight-way crosses. Three single plants of each 320 eight-way crosses were selected and propagated by single seed descent (SSD) method during five selfing generations. Finally, more than 800 eight-way tobacco MAGIC homozygous lines were constructed.

Phenotypic data analysis

Five hundred and sixty MAGIC homozygous lines (MLs) and 8 parents were planted in Zhucheng, Shandong, China (119.40_E, 36.99_N, 61 m altitude), in 2019 and 2020 (ZC2019 and ZC2020) and in Chenzhou, Hunan, China (113.01_E, 25.78_N, 190 m altitude), in 2020 (CZ2020) with a randomized complete block design with two replications. In all trials, one replication of each line contained two 10-plant rows with a row length of 10 m, row spacing of 1.2 m and plant distance of 0.5 m. Two weeks after topping, two mature middle

leaves from each plant were harvested and mixed as one sample. Three plants of each line were randomly sampled for nicotine content measurement, and the average value of three plants was used for further analysis. Nicotine content was measured using a Continuous Flow Analyzer (SAN++, Skalar Analytical B.V., Breda, the Netherlands) from ground cured leaves. Variance analysis was performed using QTL Ici mapping version 4.2.53.

Genotyping by SNP array

A total of 560 MAGIC lines and 8 parents was genotyped using the 430K tobacco SNP array with 432,362 markers at Zhengzhou Tobacco Research Institute of CNTC. Details of the genotyping and SNP calling procedures are described by Zhang et al. (2017). SNPs with no polymorphism among 8 parents, missing rate > 10% and minor allele frequency (MAF) < 0.05 were removed. Finally, a total of 93,766 SNP markers were used for subsequent studies (Table S1).

Linkage disequilibrium determination

The genome-wide LD between pairs of loci was performed using Tassel 5.0 software (Bradbury et al., 2007). The estimates

of the LD were measured using the squared allele-frequency correlations (r^2) for pairs of loci. The distance in base pairs that loci could be expected to be in LD or LD decay was computed by plotting r^2 onto physical distance using the threshold $r^2 = 0.33$ (one-third of the maximum value) as a cutoff. The LD contour plot was generated from PopLDdecay software (Zhang et al., 2019).

Genome-wide association study

All marker-trait associations were performed using TASSEL 5.0 software. A mixed linear model (MLM) was employed for GWAS analysis by incorporating a kinship matrix along with PCA. SNPs with minor allele frequency (MAF) <0.05, no polymorphisms and deletion rates higher than 10.0% were filtered out. Components of PCA were set as 5. Pedigree analysis (K) kinship method chose Centered_IBS with Max Alleles as 6. The compression level of the MLM is the optimum level, and P3D is selected for variance component estimation. Manhattan and Q-Q plots were derived from the GWAS results using the CMplot package (<https://github.com/YinLiLin/R-CMplot>) within R software (<http://www.r-project.org/>). For results with inflation, genomic control was applied to further control the confounding with population structure (Yurii et al., 2007). The Bonferroni correction method based on the FWER (Family-Wise Error Rate) standard was used to assess the significance threshold of association sites, and the significance level was determined as $p = 1.0E-07$ (0.01/93,766). When multiple loci were associated with a trait and were within a 3 Mb interval, they were considered as a single QTL.

Resequencing and data analysis

Four MAGIC parent cultivars (BH, HD, Florida301, Samsun) and one flue-cured tobacco cultivar (XHJ) were used for resequencing. Genomic DNA of the 5 tobacco accessions was randomly fragmented to 350 bp for library construction. The constructed library was sequenced on the Illumina NovaSeq 6000 platform by Novogene, Beijing, China, and 150 bp paired-end reads were generated. Furthermore, quality-controlled data were generated after removing reads with $\geq 10\%$ unidentified nucleotides (N); >10 nt aligned to the adapter, allowing $\leq 10\%$ mismatches; >50% bases having Phred quality <5 and putative PCR duplicates generated in the library construction process. After quality control of the raw data, the clean reads were mapped to the reference genome.

The candidate regions (83 Mb - 86 Mb) of chromosome 7 from five sequenced cultivars were compared, and the average coverage depth was counted using Tablet software (James Hutton Institute) with a window size of 100 kb.

DNA extraction and candidate gene cloning

Total DNA was extracted from fresh tobacco leaves using a CTAB method. The quality and quantity of total DNA was determined using a NanoDrop 2000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA). PCRs were performed with 2× Phanta Flash Master Mix (Dye Plus) and gene-specific primers according to the manufacturer's instructions. PCR products were analyzed by electrophoresis and purified using FastPure Gel DNA Extraction Mini Kit (Vazyme). The purified DNA fragment was ligated into the pEASY-T1 vector (Transgen) and then transformed into *Escherichia coli* Trans5a competent cells (Transgen) for sequencing. All primer information is listed in Table S2.

Generation of CRISPR/Cas9-mediated mutants

CRISPR/Cas9-mediated target mutagenesis was used to knock out the candidate genes, *qNIC7-1* (*NtERF170*, *NtERF199*, and *NtERF890*) and *NtERF189*, the homologous gene of *NtERF199*. The targeting sequences (listed in Table S2) and their reverse complement oligomer were chemically synthesized and annealed, and the dimers obtained were then inserted into the PHSE401 binary vector. For multigene editing, the dimers are connected with the tRNA sequence.

The constructs obtained above were individually transformed into competent *Agrobacterium tumefaciens* strain EHA105 cells using the freeze-thaw method. Leaf discs from 6-week-old XHJ were infected with *A. tumefaciens* strain EHA105 harboring the knockout vectors following the methods reported (Wang et al., 2022). The infected leaf discs were then plated onto generation medium (MS medium containing 3% (w/v) sucrose, 1 mg/L 6-benzylaminopurine, 0.15 mg/L 1-naphthaleneacetic acid, 50 mg/L cefotaxime sodium, and 8 mg/L hygromycin, pH 5.7). Hygromycin-resistant seedlings were obtained, and genomic DNA was extracted. *NtERF170*, *NtERF189*, *NtERF199* and *NtERF890* were amplified by PCR from genomic DNA, and the PCR products were Sanger sequenced to detect the mutations.

Screening of homozygous mutants and nicotine content measurement

Single-gene (M170, M199, M890) and multigene (M-2, M-3) T2 generation knockout mutants were obtained, and homozygous screening was conducted. The DNA of each individual plant was extracted. Full-length primers were used to amplify target genes, and the PCR products were purified and ligated into pEASY-T1 vectors for sequencing. At least ten clones from each sample were

sequenced. Only when all clones had identical mutations were the corresponding plants considered as homozygous mutants.

Mature middle leaves of 12-week-old plants were collected, dried in an electro-thermostatic blast oven, and ground using a tissue grinder (Jingxin Industrial Development Co., Ltd., Shanghai, China). Fifty milligrams of ground leaves were sampled for nicotine content measurement on a HITACHI Chromaster (autosampler Chromaster 5210, column oven Chromaster 5310, diode array detector Chromaster 5430, RI detector Chromaster 5450). Leaves collected from three independent plants and mixed together were considered as one replicate. Three replicates were measured for each mutant line.

Distribution analysis in germplasm resources

Different accessions were obtained from the National Medium-term GenBank of Tobacco Germplasm Resources of China. A total of 1,187 flue-cured tobacco accessions and 114 cigar accessions were planted in the field in Jimo, Shandong, China (120°44'71"E, 36°38'93"N, 22 m altitude) in 2020. The nicotine content was measured as described above. Leaf samples were collected, and genomic DNA was extracted. *NtERF189* and *NtERF199* were amplified in 1301 DNA samples using both full-length primers and short segment primers. PCR products were analyzed by electrophoresis, and those with no target amplification were considered as target gene deletions.

Results

Tobacco MAGIC population structure assessment

Eight tobacco accessions (BH, Florida301, Vam, Basma, Samsun, XHQ, TP, HD) that belong to different tobacco types were used to produce the MAGIC population (Figure 1A). The effectiveness of

random mating and its consequence on the structure of this population was evaluated using 93,766 SNP markers distributed on 24 chromosomes (Figure 1B). Population structure analysis by the neighbor-joining method indicated that no specific population structure exists among the 560 MAGIC lines (Figure 1C). Kinship analysis also showed no specific clustering pattern (Supplementary Figure 1). To determine the mapping resolution for GWAS, the average extent of LD decay was quantified (Figure 1D). Using the whole set of SNPs, the LD decay rate of the population for the entire genome was estimated to be 1,500 kb, with $r^2 = 0.33$ (one-third of the maximum value). All data indicated that the tobacco MAGIC population was suitable for GWAS analysis with no obvious population structure and moderate LD level.

Phenotypic and genotypic evaluation of the MAGIC population

Nicotine content data of 8 parents and 560 MLs were collected from three environments (2019 in Zhucheng, 2020 in Zhucheng, 2020 in Chenzhou). In this study, the nicotine content of mature middle leaves was measured, and the results are shown in Figure 2 and Supplementary Table 3. As shown in Table 1, the mean values of parents and 560 MLs were similar. However, the range of variation was more extensive in MLs, possibly due to transgressive segregation. The variation in nicotine content was determined by genotype and environment. The heritability was approximately 0.5 by plot, demonstrating that the trait was quantitatively inherited.

Genome-wide association analysis of nicotine content

GWAS was carried out using the best linear unbiased estimates (BLUES) of MLs over three replicates in a mixed

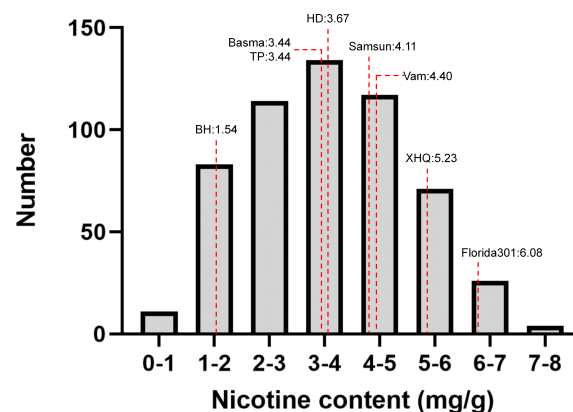


FIGURE 2
Phenotypic distribution of 8 parents and 560 MAGIC lines. Eight parents were shown as dotted lines.

TABLE 1 Phenotypic variations, variance components and heritability for nicotine content in tobacco MAGIC population in 3 environments.

	Range of Variation (mg/g)	Mean ^a (mg/g)	SD ^b	Mean squares				Heritability (%)	
				Environment	Genotype	G×E	Error	Plot level	Genotypic mean level
Parents	1.54-6.08	3.99	1.35	15.26	6.94	2.16	0.41	46.27	75.18
MLs	0.34-7.55	3.56	1.45	922.10	11.88	2.40	0.77	54.62	82.21

^aAverage nicotine content.
^bStandard deviation.

linear model (MLM) by employing Tassel 5.0 software. The quantile-quantile (Q-Q) plot showed that most of the observed p values followed a uniform distribution, but the few that were in LD with a causal polymorphism had significant p values in the tail (Supplementary Figure 2). The Manhattan plot for nicotine content was generated from TASSEL software and the R package “CMplot” (Figure 3). At genome wide significance (p value $\leq 1.0E-07$), a total of 4 QTLs ($qNIC7-1$, $qNIC7-2$, $qNIC7-3$, $qNIC5-1$) distributed on chromosome 7 and chromosome 5 were identified (Table 2). Among them, $qNIC7-1$ and $qNIC5-1$ were repeatedly detected in most environments. The other 2 QTLs ($qNIC7-2$, $qNIC7-3$) were only identified in ZC2019 and ZC2020.

The most significant locus, $qNIC7-1$, associated with nicotine content had 19 markers spanning 772 kb (Supplementary Table 4). The SNPs with the strongest association (AX-117513847, AX-117513846, AX-117632223, AX-117522763, $p = 6.85E-36$) explained 32.6% of the phenotypic variation. Based on the results of LD decay analysis and association analysis, we estimated the candidate region of $qNIC7-1$ to be 83 Mb - 86 Mb, which contained 28 SNPs and 41 annotated genes (Supplementary Table 5). In $qNIC5-1$, the SNP with the strongest association (AX-117702271, $p = 3.01E-12$) explained 9.2% of the phenotypic variation. The candidate region of $qNIC5-1$ was estimated to be 159.5 Mb-162.5 Mb, which contained 64 SNPs and 46 annotated genes (Supplementary Table 5).

Candidate gene analysis

We noticed that eight tandem ERF transcription factors were located on the region covering the strongest-associated SNPs (84 Mb - 85 Mb) in $qNIC7-1$. Resequencing data analysis showed that three of the ERFs ($NtERF170$, $NtERF199$, and $NtERF890$) were deleted in one parent cultivar, Beinhart1000-1, which had the lowest nicotine content, whereas the remaining seven showed no sequence or expression differences (Figure 4). Therefore, we focused on the three deleted ERFs. To examine whether their deletions are responsible for nicotine content reduction, we used CRISPR/Cas9 technology to generate loss-of-function mutants of three genes separately (M170, M199, M890) and collectively (M-3) in the flue-cured tobacco cultivar

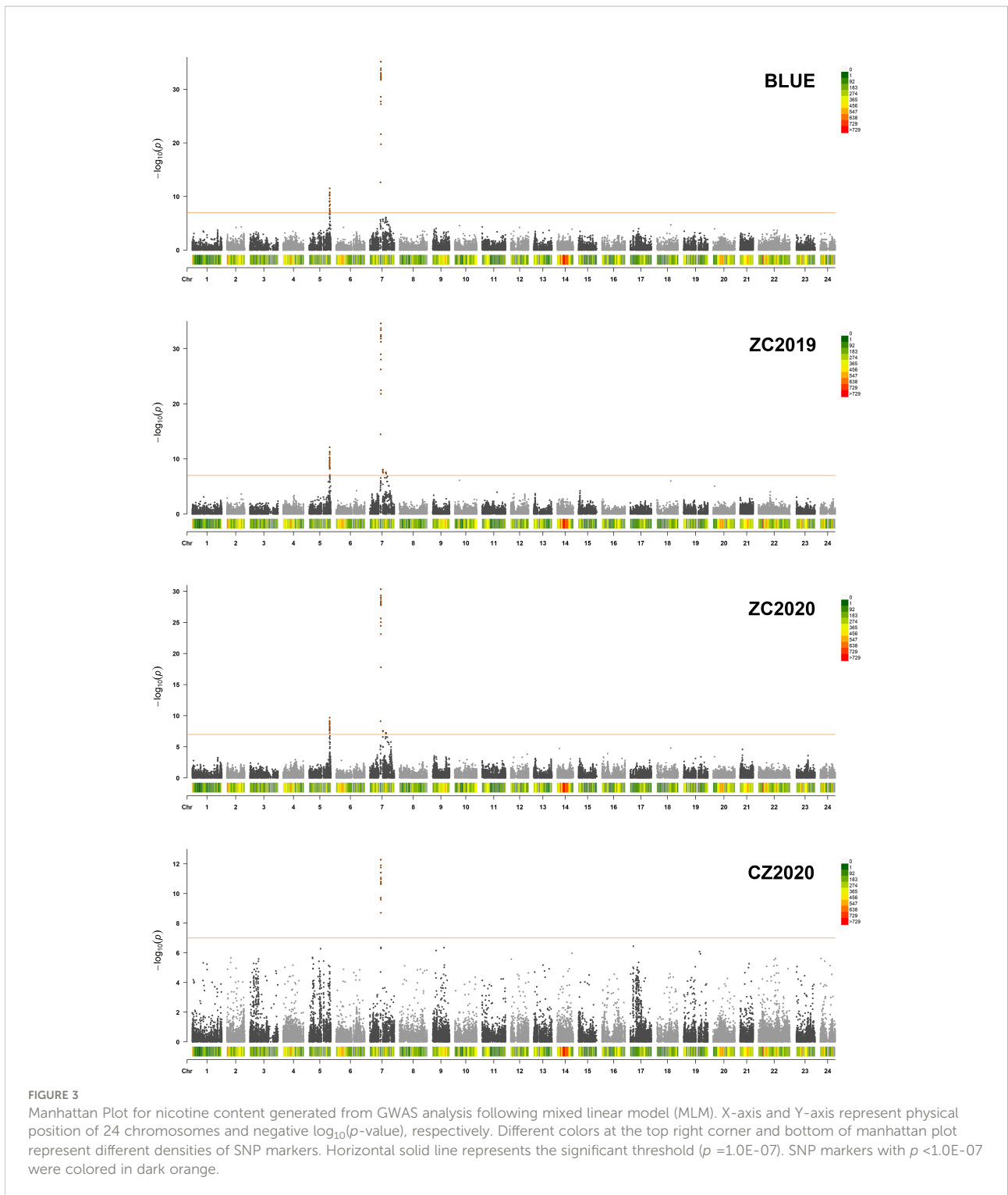
XHJ and evaluated their phenotypic effect in the T2 generation. The genotypes of wild-type (XHJ), single-gene homozygous mutants (M170, M199, M890) and multigene homozygous mutant (M-3) plants were determined by DNA sequencing. The average nicotine content of XHJ was 3.50 mg/g. The average nicotine contents of M170 and M890 were 3.37 mg/g and 3.26 mg/g, respectively, showing no obvious differences compared to the wild type. However, M199 and M-3 exhibited significant decreases with the nicotine contents of 1.19 and 0.97 mg/g, approximately 1/3 of that in wild-type XHJ, corresponding to the nicotine level in BH (1.11 mg/g) (Figure 5).

BLAST analysis showed that $NtERF189$ on chromosome 19 is homologous to $NtERF199$, sharing an identical binding domain. To investigate the effect of both genes on nicotine synthesis, we also generated a two-gene mutant (M-2) in which $NtERF189$ and $NtERF199$ were collectively knocked out (Figure 5). As expected, nicotine was barely detectable in M-2 plants (0.03 mg/g), approximately 1/100 of that in WT plants, 1/30 of BH, M199 and M-3.

The candidate region of $qNIC5-1$ contained 46 annotated genes. According to the annotation and gene function analysis, we predicted two candidate genes ($NIC5-1$ and $NIC5-2$) that might be associated with nicotine content. $NIC5-1$ was annotated as a MYB35 transcription factor and located 304 kb downstream of AX-117702271. $NIC5-2$ was located 1,188 kb downstream of AX-117702271 and homologous to the MYC2 transcription factor. The functions of the two genes in nicotine synthesis need to be further studied in the future.

Distribution analysis of mutations in germplasm resources

To explore the practical utility of the deletion mutation in $qNIC7-1$, we designed an insertion-deletion (InDel) marker (NIC7001) (forward primer: 5'-GATGTGGGTGCAACCTTTCG-3' and reverse primer: 5'-CGAACCCCTTCCTTCACATAA-3') and genotyped 560 MLs and 1301 tobacco accessions. The InDel marker NIC7001 locus has two types: deletion type and nondeletion type (839 bp amplification). The two types were compared using one-way ANOVA using IBM SPSS Statistics software version 23 (IBM, New York, USA). In the MAGIC



population, the lines of the nondeletion type had a mean nicotine content of 3.95 mg/g, which was significantly higher than the mean nicotine content of the deletion type lines (1.77 mg/g). In 1301 tobacco cultivars, which included 1187 flue-cured

tobacco accessions and 114 cigar accessions, only 11 cultivars (8 flue-cured tobacco types and 3 cigar tobacco types) were detected as deletion types (Table 3). Five out of the 11 cultivars, together with six nondeletion-type cultivars, were

TABLE 2 Summary of QTLs associated with nicotine content in tobacco across different environments.

QTL	SNP	Position	ZC2019 ^a		ZC2020 ^a		CZ2020 ^a		BLUE ^b	
			<i>p</i>	effect	<i>p</i>	effect	<i>p</i>	effect	<i>p</i>	effect
<i>qNIC7-1</i>	AX-117513847	84,056,268	2.59E-35	2.57	4.49E-31	1.90	5.21E-13	1.97	6.85E-36	2.09
	AX-117513846	84,098,896	2.59E-35	2.57	4.49E-31	1.90	5.21E-13	1.97	6.85E-36	2.09
	AX-117632223	84,282,622	2.59E-35	2.57	4.49E-31	1.90	5.21E-13	1.97	6.85E-36	2.09
	AX-117522763	84,740,457	2.59E-35	2.57	4.49E-31	1.90	5.21E-13	1.97	6.85E-36	2.09
<i>qNIC7-2</i>	AX-117702600	99,815,696	8.95E-09	-0.05	2.66E-08	-0.27	/	/	/	/
<i>qNIC7-3</i>	AX-117503751	121,872,699	2.85E-08	1.54	5.77E-08	1.16	/	/	/	/
	AX-117538002	124,443,130	2.85E-08	1.54	5.77E-08	1.16	/	/	/	/
<i>qNIC5-1</i>	AX-117702271	160,770,794	4.82E-12	1.16	2.10E-10	0.77	/	/	3.01E-12	0.92

^aZC2019, ZC2020 and CZ2020 refer to Zhucheng in 2019, Zhucheng in 2020 and Chenzhou in 2020.
^bThe best linear unbiased estimates

phenotyped. In accordance with the results in MAGIC lines, the nicotine content of deletion type cultivars was significantly lower than that of nondeletion type cultivars (Figure 6).

The distribution of *NtERF189* (NIC2) deletion was also investigated using the InDel marker NIC7002 (forward primer: 5'-GCGTGCAATAGGGCAAAGC-3' and reverse primer: 5'-TGAAAGCTTCCTTCCTTTTACACA-3') in 1301 tobacco germplasm cultivars (Table 3). The deletion of *NtERF189* was detected in 14 flue-cured tobacco accessions (1.18%) and 15 cigar accessions (13.16%), with a total deletion rate of 2.23%. *NtERF199* was deleted in only 8 flue-cured tobacco accessions (0.67%) and 3 cigar accessions (2.63%), with a total deletion rate of 0.85%. No accession was identified that carried deletions in both genes.

Discussion

The MAGIC population exhibits more power and potential for QTL mapping

To identify genetic variation controlling complex traits variation, many biparental populations have been widely used for QTL mapping in tobacco (Drake-Stowe et al., 2017; Sun et al., 2018; Agacka-Mołodoch et al., 2021). However, the main limitations of biparental population for QTL mapping are the low genetic diversity of the mapping population and mapping resolution. In MAGIC population, multiple founders are intercrossed several times in a well-defined order to combine the genetic material of all the founders in a single line (Cavanagh et al., 2008). Therefore, compared to biparental mapping population, MAGIC population contains more recombination events and genetic diversity, which enhances mapping resolution and the ability to analyze several alleles simultaneously. In this study, eight founders in the MAGIC population enriched population with higher allelic diversity compared to those derived from the traditional biparental

populations. Furthermore, we narrowed down the target region, and *NtERF199* was considered exactly as the candidate gene by MAGIC population. In addition, the distribution of *NtERF199* in tobacco germplasm resources was analyzed. The deletion of *NtERF199* was considered as a rare allelic variation (RAV) with a deletion rate of 0.85%. Sun et al. (2017) performed GWAS using 219 flue-cured tobacco accessions and identified 2 QTLs associated with nicotine content by GWAS. However, the most major-effect QTL (*qNIC7-1*) in this study was not identified using the 219 flue-cured tobacco accessions, probably due to the filtration of RAV. Because the crossover technique elevates the frequency of all parental alleles to be almost equal, MAGIC population have higher power to identify low frequency alleles and can better evaluate allelic effects between founders than genome-wide association panels (Dell'Acqua et al., 2015; Scott et al., 2020). Another main advantage of MAGIC population over association panels is the lack of an underlying unknown structure, which enhances the risk of detecting false-positives. Therefore, MAGIC population represents an intermediate resources between biparental crosses and association panels regarding the number of traits that can be investigated, resolution, allele diversity and population structure (Rakshit et al., 2012; Pascual et al., 2016). We believe that this tobacco MAGIC population provides a good resource for studying the genetic basis of complex traits in tobacco.

GWAS and identification of candidate genes for nicotine content

GWAS is a forward genetic approach to identify underlying causal genes, mutations and putative functional markers that affect complex quantitative traits. In this study, we identified four QTLs distributed on two chromosomes associated with nicotine content. The candidate gene *NtERF199* on *qNIC7-1* and its homologous gene *NtERF189* were cloned and experimentally

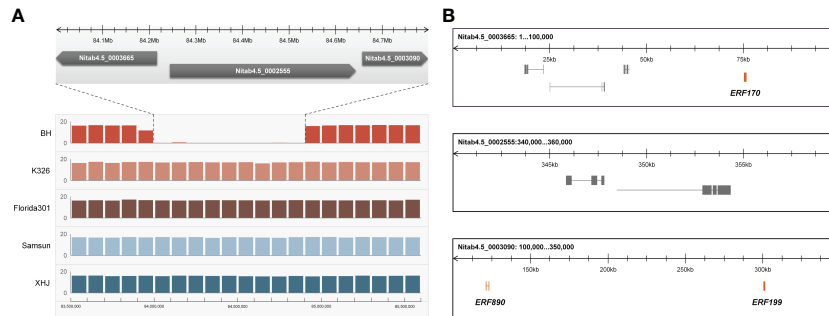


FIGURE 4 Candidate gene analysis. (A) Resequencing data of 5 tobacco accessions (BH, K326, Florida301, Samsun, XHJ) revealed a fragment deletion in BH. X-axis represent position of Chr07 with the window size of 100kb. Y-axis represent average coverage depth of resequencing data. The deletion fragment was mapped to 3 scaffolds (Nitab4.5_0003665, Nitab4.5_0002555, Nitab4.5_0003090) of K326. (B) The 3 scaffolds of K326 corresponding to the deleted fragment comprises of three ERFs (*NtERF170*, *NtERF199* and *NtERF870*).

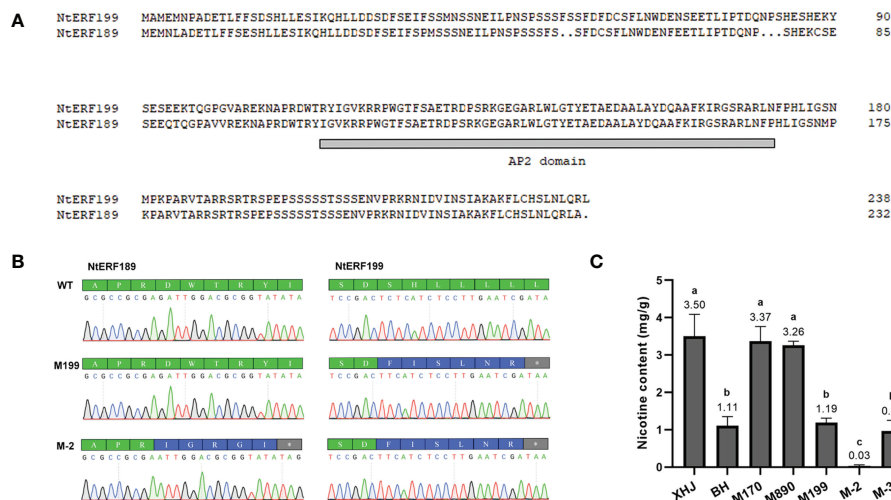
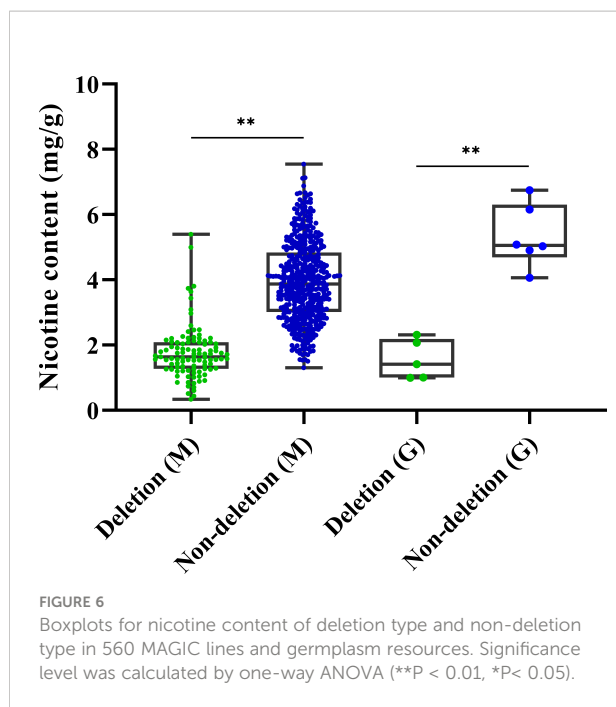


FIGURE 5 Validation of the association between two ERFs and nicotine content. (A) Alignment of *NtERF199* and *NtERF189*. The AP2 domain is indicated by a gray line. (B) Mutation of *NtERF189* in M-2 and *NtERF199* in M199 and M-2. Normal amino acids are marked in green. Mutated amino acids are marked in blue and stop codon is marked in grey. (C) Nicotine content of wildtype (XHJ, BH) and T2 homozygous mutants (M170, M890, M199, M-2, M-3).

TABLE 3 Distribution of deletion mutation of *NtERF189* and *NtERF199* in germplasm resources.

	Accessions Tested	<i>NtERF199</i> deleted	<i>NtERF189</i> deleted	Both deleted
Flue-cured tobacco accessions	1187	8 (0.67%)	14 (1.18%)	0
Cigar accessions	114	3 (2.63%)	15 (13.16%)	0
Total	1301	11 (0.85%)	29 (2.23%)	0



verified through loss-of-function mutant analysis. The results are consistent with previously reported studies. Shoji et al. (2010) identified that the *NIC2* locus comprises a cluster of ERFs (including *NtERF189*) that regulate nicotine biosynthesis in tobacco. Qin et al., (2021) validated *NtERF199* as the *NIC1* gene. The detection of *qNIC7-1* highlighted the strength of using tobacco MAGIC population in GWAS.

As in the case of *NIC2* locus in tobacco, eight ERFs are clustered together in the *qNIC7-1* locus. However, both the previous study by Qin et al., (2021) and our investigation confirmed that *NtERF199* is the only causal gene for nicotine biosynthesis within the *qNIC7-1* locus. Same case also exists in tomatoes, with *GAME9* locates within an ERF cluster and works as the only functional regulator of steroidal glycoalkaloid biosynthesis (Cardenas et al., 2016). These results suggest potential functional redundancy and divergence of clustered ERFs.

Allopolyploid plants usually display higher adaptability than their diploids ancestors since their preexisting homeologous loci from the progenitors show a predominantly additive gene expression pattern (Fu et al., 2016). As a natural allopolyploid plant, *N. tabacum* was formed by hybridization of *Nicotiana sylvestris* (S-genome) and *Nicotiana tomentosiformis* (T-genome) (Clarkson et al., 2017). Blast analysis and functional studies have demonstrated that *NtERF199* and *NtERF189* are a pair of homologous genes originated from *N. sylvestris* and *N. tomentosiformis*, respectively, and showed dose-dependent effects on nicotine levels (Shoji et al., 2010; Qin et al., 2021). It has also been proven through transcriptome analysis that most homologous genes from the S- and T-genomes of tobacco have

been maintained and are expressed in a similar way (Bombarely et al., 2012; Edwards et al., 2017). Therefore, the effect of homologous genes must be taken into consideration in the study of gene function in allotetraploid plants, especially *N. tabacum*.

We also identified a significant QTL on chromosome 5 (*qNIC5-1*), which explained 9.6% of the phenotypic variation. Among the 46 annotated genes in *qNIC5-1*, we predicted two candidate genes (*NIC5-1* and *NIC5-2*) likely associated with nicotine content based on the annotation and gene function analysis. *NIC5-1* was annotated as a MYB35 transcription factor. MYB35 has been reported to regulate methyl jasmonate (MeJA) and wound-responsive G10H-1 in *Bacopa monnieri* (Jeena et al., 2021). Nicotine biosynthesis in tobacco is highly regulated by jasmonic acid (JA) (Baldwin et al., 1996; Shoji et al., 2008). Therefore, MYB35 might participate in nicotine synthesis by regulating JA-responsive and nicotine biosynthesis-associated genes. *NIC5-2* was annotated as a MYC2 transcription factor, which has been described to control nicotine biosynthesis genes by directly binding the G-box in the target promoters and upregulating the *NIC2*-locus ERF genes (Shoji and Hashimoto, 2011; Zhang et al., 2012; Sui et al., 2021).

Development of functional markers and distribution of allelic variation

To investigate the distribution of *NtERF189* and *NtERF199* deletion, we manually designed two InDel markers (NIC7001 and NIC7002) and screened a total of 1301 cultivated tobacco accessions, including 1187 flue-cured tobacco accessions and 114 cigar tobacco accessions. In flue-cured tobacco, the deletion mutation rates of *NtERF189* and *NtERF199* were 1.18% and 0.67%, respectively, while in cigar tobacco, they reached 13.16% and 2.63%, respectively. The results demonstrated that the two ERF deletion mutations displayed higher frequency in cigar tobacco. In fact, the low-nicotine trait was initially found in strains of Cuban cigar tobacco varieties and subsequently introduced into cigarette varieties through a series of backcrosses (Valleau, 1949). This might explain the higher mutation frequency in cigar tobacco. Overall, the deletion mutation of *NtERF189* and *NtERF199* was detected in only 2.23% and 0.85% of the tested tobacco cultivar, demonstrating them as rare allelic variations. Deletion of both genes was detected in no accessions. This may be attributed to the importance of nicotine in resisting biological stresses, and those with both ERFs mutated have already been naturally eliminated.

Moreover, we validated the effectiveness of NIC7001 using a panel of 11 tobacco cultivars (5 deletion types and 6 nondeletion types). The results provide promising evidence of InDel marker application for future breeding programs of low nicotine tobacco cultivars.

Conclusion

This work reports the first MAGIC population in tobacco, constructed from eight common breeding lines. By using the MAGIC population with negligible population structure coupled with high-density SNP markers, GWAS was performed and identified four QTLs associated with the low nicotine trait. The candidate gene *NtERF199* in *qNIC7-1* and its homologous gene *NtERF189* were analyzed and verified, which proved the reliability of the GWAS results in the MAGIC population. Candidate genes in *qNIC5-1* were also analyzed and need to be further studied. Overall, the tobacco MAGIC population provides for the identification of QTLs associated with other complicated important traits.

Data availability statement

The datasets presented in this study can be found in online repositories. The name of the repository and accession number(s) can be found below: NCBI Sequence Reads Archive (SRA) database (Accession number: PRJNA904563).

Author contributions

GY, KS, WY, ZJ, LW, and FW performed the experiment. GY, LC and HS analyzed the experimental data. CJ, DL and HM participated in the field trails. YW, AY and LC designed the experiment. GY and LC wrote this manuscript. All authors contributed to the article and approved the submitted version.

References

- Agacka-Mołdoch, M., Rehman-Arif, M. A., Lohwasser, U., Doroszewska, T., Lewis, R. S., and Börner, A. (2021). QTL analysis of seed germination traits in tobacco (*Nicotiana tabacum* L.). *J. Appl. Genet.* 62, 441–444. doi: 10.1007/s13353-021-00623-6
- Arrones, A., Vilanova, S., Plazas, M., Mangino, G., Pascual, L., Díez, M. J., et al. (2020). The dawn of the age of multi-parent MAGIC populations in plant breeding: Novel powerful next-generation resources for genetic analysis and selection of recombinant elite material. *Biology* 9, 229. doi: 10.3390/biology9080229
- Baldwin, I. T., Schmelz, E. A., and Zhang, Z. P. (1996). Effects of octadecanoid metabolites and inhibitors on induced nicotine accumulation in *Nicotiana sylvestris*. *J. Chem. Ecol.* 22, 61–74. doi: 10.1007/BF02040200
- Bombarely, A., Edwards, K. D., Sanchez-Tamburrino, J., and Mueller, L. A. (2012). Deciphering the complex leaf transcriptome of the allotetraploid species *Nicotiana tabacum*: A phylogenomic perspective. *BMC Genomics* 13, 406. doi: 10.1186/1471-2164-13-406
- Bossa-Castro, A. M., Tekete, C., Raghavan, C., Delorean, E. E., Dereeper, A., Dagno, K., et al. (2018). Allelic variation for broad-spectrum resistance and susceptibility to bacterial pathogens identified in a rice MAGIC population. *Plant Biotechnol. J.* 16, 1559–1568. doi: 10.1111/pbi.12895
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., Buckler, E. S., et al. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Funding

This work was supported by grants from the Agricultural Science and Technology Innovation Program (ASTIP-TRIC01).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.1086950/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Kinship matrix among the 560 MAGIC lines of tobacco MAGIC population using SNP markers.

SUPPLEMENTARY FIGURE 2

Q-Q plots of nicotine content across different environments.

- Cardenas, P. D., Sonawane, P. D., Pollier, J., Vanden Bossche, R., Dewangan, V., Weithorn, E., et al. (2016). GAME9 regulates the biosynthesis of steroidal alkaloids and upstream isoprenoids in the plant mevalonate pathway. *Nat. Commun.* 7, 10654. doi: 10.1038/ncomms10654

Cavanagh, C., Morell, M., Mackay, I., and Powell, W. (2008). From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Curr. Opin. Plant Biol.* 11, 215–221. doi: 10.1016/j.pbi.2008.01.002

Chen, X., Fu, S., Zhang, P., Gu, Z., Liu, J., Qian, Q., et al. (2013). Proteomic analysis of a disease-resistance-enhanced lesion mimic mutant spotted leaf 5 in rice. *Rice (N.Y)* 6, 1. doi: 10.1186/1939-8433-6-1

Clarkson, J. J., Dodsworth, S., and Chase, M. W. (2017). Time-calibrated phylogenetic trees establish a lag between polyploidisation and diversification in *Nicotiana* (Solanaceae). *Plant Syst. Evol.* 303, 1001–1012. doi: 10.1007/s00606-017-1416-9

Dell'Acqua, M., Gatti, D. M., Pea, G., Cattonaro, F., Coppens, F., Magris, G., et al. (2015). Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in *zea mays*. *Genome Biol.* 16, 167. doi: 10.1186/s13059-015-0716-z

Doerge, R. W. (2002). Mapping and analysis of quantitative trait loci in experimental populations. *Nat. Rev. Genet.* 3, 43–52. doi: 10.1038/nrg703

Drake-Stowe, K., Bakaher, N., Goepfert, S., Philippon, B., Mark, R., Peterson, P., et al. (2017). Multiple disease resistance loci affect soilborne disease resistance in

- tobacco (*Nicotiana tabacum*). *Phytopathology* 107, 1055–1061. doi: 10.1094/PHYTO-03-17-0118-R
- Edwards, K. D., Fernandez-Pozo, N., Drake-Stowe, K., Humphry, M., Evans, A. D., Bombarely, A., et al. (2017). A reference genome for *Nicotiana glauca* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. *BMC Genomics* 18, 448. doi: 10.1186/s12864-017-3791-6
- Falconer, D. S., and Mackay, T. F. C. (1996). *Introduction to quantitative genetics*, 4th ed (Longman: Harlow, UK).
- Fu, D., Mason, A. S., Xiao, M., and Yan, H. (2016). Effects of genome structure variation, homeologous genes and repetitive DNA on polyploid crop research in the age of genomics. *Plant Sci.* 242, 37–46. doi: 10.1016/j.plantsci.2015.09.017
- Hibi, N., Fujita, T., Hatano, M., Hashimoto, T., and Yamada, Y. (1992). Putrescine n-methyltransferase in cultured roots of *hyoscyamus albus*. *Plant Physiol.* 100, 826–835. doi: 10.1104/pp.100.2.826
- Hibi, N., Higashiguchi, S., Hashimoto, T., and Yamada, Y. (1994). Gene expression in tobacco low-nicotine mutants. *Plant Cell* 6, 723–735. doi: 10.1105/tpc.6.5.723
- Holland, J. B. (2007). Genetic architecture of complex traits in plants. *Curr. Opin. Plant Biol.* 10, 156–161. doi: 10.1016/j.pbi.2007.01.003
- Huang, B. E., Verbyla, K. L., Verbyla, A. P., Raghavan, C., Singh, V. K., Gaur, P., et al. (2015). MAGIC populations in crops: current status and future prospects. *Theor. Appl. Genet.* 128, 999–1017. doi: 10.1007/s00122-015-2506-0
- Jaiswal, V., Gupta, S., Gahlaut, V., Muthamilarasan, M., Bandyopadhyay, T., Ramchiari, N., et al. (2019). Genome-wide association study of major agronomic traits in foxtail millet (*Setaria italica* L.) using ddRAD sequencing. *Sci. Rep.* 9, 5020. doi: 10.1038/s41598-019-41602-6
- Jeena, G. S., Kumar, S., and Shukla, R. K. (2021). Characterization of MYB35 regulated methyl jasmonate and wound responsive geraniol 10-hydroxylase-1 gene from *Bacopa monnieri*. *Planta* 253, 89. doi: 10.1007/s00425-021-03614-3
- Kajikawa, M., Hirai, N., and Hashimoto, T. (2009). A PIP-family protein is required for biosynthesis of tobacco alkaloids. *Plant Mol. Biol.* 69, 287–298. doi: 10.1007/s11103-008-9424-3
- Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., et al. (2009). A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genet.* 5, e1000551. doi: 10.1371/journal.pgen.1000551
- Legg, P. D., and Collins, G. B. (1971). Inheritance of percent total alkaloids in *Nicotiana glauca* L. II. genetic effects of two loci in Burley21 × LA burley 21 populations. *Can. J. Genet. Cytol.* 13, 287–291. doi: 10.1139/g71-047
- Mackay, I., and Powell, W. (2007). Methods for linkage disequilibrium mapping in crops. *Trends Plant Sci.* 12, 57–63. doi: 10.1016/j.tplants.2006.12.001
- Mitchell-Olds, T. (2010). Complex-trait analysis in plants. *Genome Biol.* 11, 113. doi: 10.1186/gb-2010-11-4-113
- Ongom, P. O., and Ejeta, G. (2018). Mating design and genetic structure of a multiparent advanced generation intercross (MAGIC) population of sorghum (*Sorghum bicolor* (L.) Moench). *G3 (Bethesda)* 8, 331–341. doi: 10.1534/g3.117.300248
- Pascual, L., Albert, E., Sauvage, C., Duangjit, J., Bouchet, J. P., Bitton, F., et al. (2016). Dissecting quantitative trait variation in the resequencing era: Complementarity of bi-parental, multi-parental and association panels. *Plant Sci.* 242, 120–130. doi: 10.1016/j.plantsci.2015.06.017
- Pascual, L., Desplat, N., Huang, B. E., Desgroux, A., Bruguier, L., Bouchet, J. P., et al. (2015). Potential of a tomato MAGIC population to decipher the genetic control of quantitative traits and detect causal variants in the resequencing era. *Plant Biotechnol. J.* 13, 565–577. doi: 10.1111/pbi.12282
- Qin, Q., Humphry, M., Gilles, T., Fisher, A., Patra, B., Singh, S. K., et al. (2021). NIC1 cloning and gene editing generates low-nicotine tobacco plants. *Plant Biotechnol. J.* 19 (11), 2150–2152. doi: 10.1111/pbi.13694
- Rakshit, S., Rakshit, A., and Patil, J. V. (2012). Multiparent intercross populations in analysis of quantitative traits. *J. Genet.* 91, 111–117. doi: 10.1007/s12041-012-0144-8
- Remington, D. L., Thornsberry, J. M., Matsuoka, Y., Wilson, L. M., Whitt, S. R., Doebley, J., et al. (2001). Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc. Natl. Acad. Sci.* 98, 11479–11484. doi: 10.1073/pnas.201394398
- Rollar, S., Serfling, A., Geyer, M., Hartl, L., Mohler, V., and Ordon, F. (2021). QTL mapping of adult plant and seedling resistance to leaf rust (*Puccinia triticina* Erikss.) in a multiparent advanced generation intercross (MAGIC) wheat population. *Theor. Appl. Genet.* 134, 37–51. doi: 10.1007/s00122-020-03657-2
- Ryan, S. M., Cane, K. A., DeBoer, K. D., Sinclair, S. J., Brimblecombe, R., and Hamill, J. D. (2012). Structure and expression of the quinolinate phosphoribosyltransferase (QPT) gene family in *Nicotiana glauca*. *Plant Sci.* 188–189, 102–110. doi: 10.1016/j.plantsci.2012.02.008
- Scott, M. F., Ladejobi, O., Amer, S., Bentley, A. R., Biernaskie, J., Boden, S. A., et al. (2020). Multi-parent populations in crops: A toolbox integrating genomics and genetic mapping with breeding. *Heredity* 125, 396–416. doi: 10.1038/s41437-020-0336-6
- Shoji, T., and Hashimoto, T. (2011). Tobacco MYC2 regulates jasmonate-inducible nicotine biosynthesis genes directly and by way of the NIC2-locus ERF genes. *Plant Cell Physiol.* 52, 1117–1130. doi: 10.1093/pcp/pcr063
- Shoji, T., Kajikawa, M., and Hashimoto, T. (2010). Clustered transcription factor genes regulate nicotine biosynthesis in tobacco. *Plant Cell* 22, 3390–3409. doi: 10.1105/tpc.110.078543
- Shoji, T., Ogawa, T., and Hashimoto, T. (2008). Jasmonate-induced nicotine formation in tobacco is mediated by tobacco COI1 and JAZ genes. *Plant Cell Physiol.* 49, 1003–1012. doi: 10.1093/pcp/pcn077
- Sinclair, S. J., Murphy, K. J., Birch, C. D., and Hamill, J. D. (2000). Molecular characterization of quinolinate phosphoribosyltransferase (QPRTase) in *Nicotiana glauca*. *Plant Mol. Biol.* 44, 603–617. doi: 10.1023/a:1026590521318
- Sui, X., He, X., Song, Z., Gao, Y., Zhao, L., Jiao, F., et al. (2021). The gene NtMYC2a acts as a 'master switch' in the regulation of JA-induced nicotine accumulation in tobacco. *Plant Biol. (Stuttg)* 23, 317–326. doi: 10.1111/plb.13223
- Sun, M., Cheng, L., Jiang, C., Zhu, C., Ren, M., Zhang, Y., et al. (2018). Identification of a major QTL affecting resistance to brown spot in tobacco (*Nicotiana glauca* L.) via linkage and association mapping methods. *Euphytica* 214, 195. doi: 10.1007/s10681-018-2244-x
- Sun, X., Wen, L., Yin, P., Luo, C., Fan, W., and Ren, M. (2017). Genome-wide association study of the nicotine content in flue-cured tobacco. *Acta Tabacaria Sinica.* 23, 102–113.
- Thyssen, G. N., Jenkins, J. N., McCarty, J. C., Zeng, L., Campbell, B. T., Delhom, C. D., et al. (2019). Whole genome sequencing of a MAGIC population identified genomic loci and candidate genes for major fiber quality traits in upland cotton (*Gossypium hirsutum* L.). *Theor. Appl. Genet.* 132, 989–999. doi: 10.1007/s00122-018-3254-8
- Valdar, W., Flint, J., and Mott, R. (2006). Simulating the collaborative cross: power of quantitative trait loci detection and mapping resolution in large sets of recombinant inbred strains of mice. *Genetics* 172, 1783–1797. doi: 10.1534/genetics.104.039313
- Valleau, W. D. (1949). Breeding low-nicotine tobacco. *J. Agric. Res.* 78, 171–181.
- Visscher, P. M., Brown, M. A., McCarthy, M. I., and Yang, J. (2012). Five years of GWAS discovery. *Am. J. Hum. Genet.* 90, 7–24. doi: 10.1016/j.ajhg.2011.11.029
- Wang, Z., Yan, X., Zhang, H., Meng, Y., Pan, Y., and Cui, H. (2022). NtCycB2 negatively regulates tobacco glandular trichome formation, exudate accumulation, and aphid resistance. *Plant Mol. Biol.* 108, 65–76. doi: 10.1007/s11103-021-01222-z
- Yamamoto, E., Iwata, H., Tanabata, T., Mizobuchi, R., Yonemaru, J., Yamamoto, T., et al. (2014). Effect of advanced intercrossing on genome structure and on the power to detect linked quantitative trait loci in a multi-parent population: A simulation study in rice. *BMC Genet.* 15, 50. doi: 10.1186/1471-2156-15-50
- Yurii, S. A., Stephan, R., Aaron, I., and Cornelia, M. (2007). GenABEL: An R library for genome-wide association analysis. *Bioinformatics* 23, 1294–1296. doi: 10.1093/bioinformatics/btm108
- Zhang, H. B., Bokowiec, M. T., Rushton, P. J., Han, S. C., and Timko, M. P. (2012). Tobacco transcription factors NtMYC2a and NtMYC2b form nuclear complexes with the NtJAZ1 repressor and regulate multiple jasmonate-inducible steps in nicotine biosynthesis. *Mol. Plant* 5, 73–84. doi: 10.1093/mp/ssr056
- Zhang, C., Dong, S. S., Xu, J. Y., He, W. M., and Yang, T. L. (2019). PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* 35, 1786–1788. doi: 10.1093/bioinformatics/bty875
- Zhang, J. F., Luo, Z. P., He, S. B., Jin, J. J., Li, Z. F., Xu, Y. L., et al. (2017). Genetic diversities of 24 tobacco cultivars analyzed by SNP. *Tobacco Sci. Technol.* 50, 1–8.